ERDC/EL TR-01-14

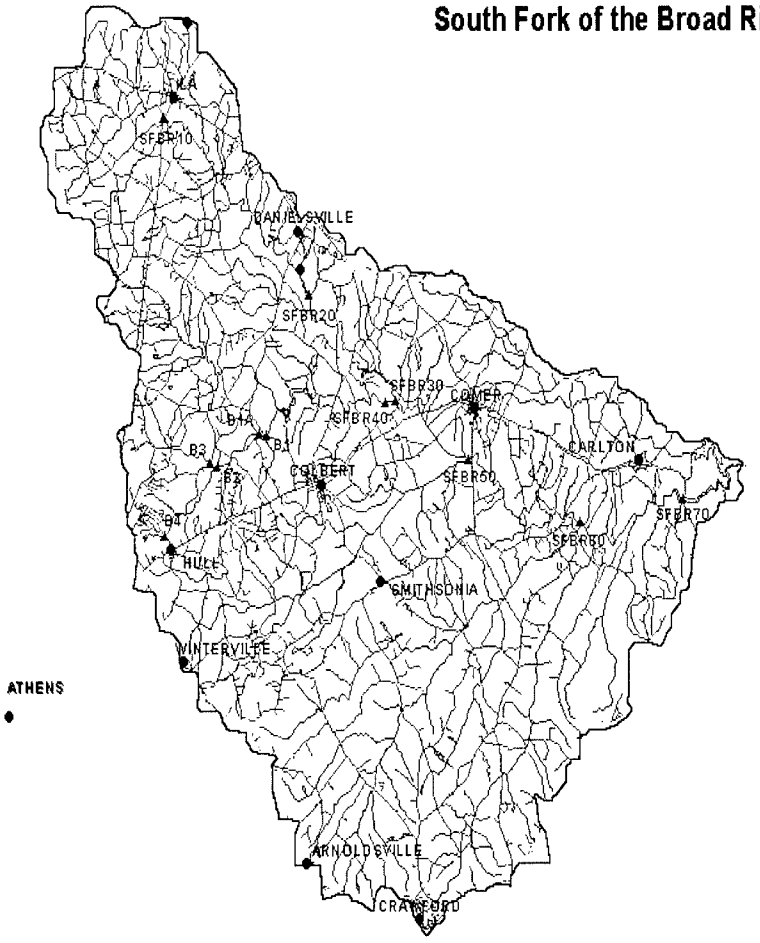
Environmental Laboratory

US Army Corps
of Engineers®
Engineer Research and
Development Center

Integrated Electronic Data Repository to Support the South Fork of the Broad River Research Program

Linda Peyman Dove and Steve Bong

August 2001



The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products.

The findings of this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

PRINTED ON RECYCLED PAPER

ERDC/EL TR-01-14
August 2001

Integrated Electronic Data Repository to Support the South Fork of the Broad River Research Program

by Linda Peyman Dove

Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Steve Bong

Jaya Corporation
4900 University Square, Suite 30
Huntsville, AL 35816

Final report

Prepared for Ecosystems Research Division
National Exposure Research Laboratory
U.S. Environmental Protection Agency
960 College Station Road, Athens, GA 30605

Contents

Preface

This report was prepared for the Ecosystems Research Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency (USEPA), Athens, GA. It is the summary of the first part of the Integrated Electronic Data Repository (IEDR) to support all phases of the multiyear Technology Required for Alternative Analyses for a Changing Environment (TRACE) Program. The project was developed by the Environmental Laboratory (EL), U.S. Army Engineer Research and Development Center (ERDC), Vicksburg, MS, and Jaya Corporation, Huntsville, AL. Ms. Brenda Kitchens, USEPA, has served as the point of contact between ERDC and the USEPA.

Authors of this report were Dr. Linda Peyman Dove, Environmental Systems Branch (EE-C), EL, ERDC, and Mr. Steve Bong, Jaya Corporation. The study was under the general supervision of Dr. Rose Kress, EE-C, and Mr. Harold W. West, Chief, EE-C; Dr. Dave Tazik, Chief, Ecosystem Evaluation and Engineering Division; and Dr. Edwin A. Theriot, Acting Director, EL.

At the time of publication of this report, Dr. James R. Houston was Director of ERDC, and COL John W. Morris III, EN, was Commander and Executive Director.

This report should be cited as follows:

> Dove, L. P., and Bong, S. (2001). "Integrated electronic data to support the South Fork of the Broad River research program," ERDC/EL TR-01-14, U.S. Army Engineer Research and Development Center, Vicksburg, MS.

1 Background

The U.S. Army Engineer Research and Development Center (ERDC), Vicksburg, MS, is currently developing a design and implementation plan for an Integrated Electronic Data Repository (IEDR) to support all phases of the multiyear Technology Required for Alternative Analyses for a Changing Environment (TRACE) Program. Specifically, the repository is needed to support the Total Maximum Daily Load (TMDL) field research projects established in the South Fork of the Broad River watershed. The purpose of the repository is to archive and disseminate various types of data associated with the project to support environmental monitoring and modeling. The modeling activities include those such as sediment transport, watershed analyses, and environmental analyses. These models require data gathered ultimately from various data files submitted and cataloged in an orderly fashion within the repository.

This report summarizes the design requirements and considerations used to develop a prototype web based repository site for the TMDL research project. It is presented in two parts: design requirements and considerations and prototype web-based repository.

2 Design Requirements and Considerations

The following summarizes the design requirements either discussed at the 24 August 2000 meeting between the U.S. Environmental Protection Agency (USEPA), ERDC, and Jaya Corporation, or developed in response to this meeting. These design requirements have been used to develop a prototype web-based repository. The overall concept of the IEDR is presented in Figure 1.

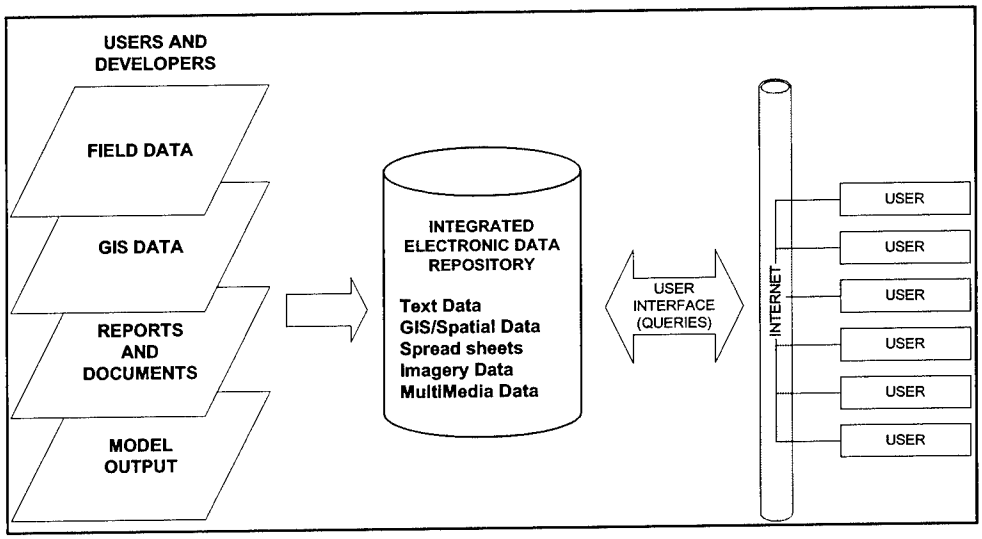


Figure 1. Integrated electronic data repository concept

a. *Architecture*: The server architecture to support the repository may be centralized or decentralized (distributed). The implemented design should be one that best serves the anticipated data and data warehousing requests. The system will be designed initially as an intranet application accessible to only members of the research effort or project. It is estimated that the number of users will not be greater than 10 and the number of concurrent users will not be greater than 2. If applicable, the system will eventually be migrated to an Internet application.

b. *Hardware and Software*: The system will be designed and implemented in the Microsoft NT 4.0 operating system. The IEDR will be

implemented on a NT 4.0 Windows Server with the following minimum hardware requirements: processor speed of 600 MHz, memory of 256Mb, tape backup capability. The repository will be designed for the above NT Windows platform running Internet Information Server (IIS), Microsoft FrontPage 2000, Microsoft Access, Active Server Page (ASP) upload, and Persits server software.

Repository users will be required to use a personal computer running Microsoft Explorer 4.0, Netscape 4.75, or greater. To maximize the repository's functionality, users will be required to have e-mail access and an e-mail program available on their computer.

c. *Data Sources*: The system will be capable of storing and retrieving files in several formats, such as text, spatial, tabular/spreadsheet, and video/ multimedia files. Major sources of data include: collected field data, laboratory analysis, research analysis/reports, and spatial/Geographic Information System (GIS)/imagery data. Examples might include:

(1) Field Data

- Stream data for the six study sites
- Samples/data to be collected before and during each storm event for all monitoring sites
- Stream flow data collected between storm events
- Weather station data
- Additional rain gauges
- Stream cross sections
- Data from real time gauging stations at Carlton site
- Rating curves for each sampling station

(2) Laboratory Analysis

- Suspended sediment data
- Bedload sediment data
- Nutrient data
- Bacteria data
- Instrumentation data

(3) Research Analysis/Reports

- Correlation of pollutants in groundwater from sediment data analyses
- Transport mechanisms of pollutants in streams
- Riparian zone and restoration

(4) Spatial/GIS/Imagery

- Digital Orthophoto Quadrangle (DOQ)
- Satellite Imagery
- GIS Data Layers

d. *Query Capability:* Server capability will support search and retrieval of data stored in a directory tree structure.

Data files within the data file repository should be organized by themes. Data files stored in the repository will generally fall into one of the following themes (categories). These themes represent an initial set and may be expanded as the repository matures.

- Surface Hydrology Data such as flow and discharge data
- Water Quality (nutrients) Data such as nitrate, phosphate content data
- Watershed Characterization Data such as site characterization data
- Sediment Data
- Documentation or Reports
- Pathogens Data
- Stream Data
- Meteorological/Weather Data

Although data will be classified into one of the above themes, it will normally be cross referenced to a specific date, monitoring site, keyword, contributor, and/or data type. This will provide sufficient cross-referencing to accommodate the user's search requirements.

e. *Quality Control*: The system will have a mechanism that will check all incoming data for viruses. The system is not required to automatically check the accuracy of data content being submitted. This responsibility rests with the contributor, author or originator of the data. Each incoming data file will be checked to ensure that the submitted data include a description and metadata file.

f. *Security*: Several levels of security layers for accessing the site and data files must be provided. The initial system will be restricted to use by a select USEPA user group. When the system is fully implemented, consideration will be given for limited general public access.

3 Prototype Web-Based Repository Design

Conceptual Design

A prototype web-based repository for the TMDL project at South Fork of the Broad River was developed and can be viewed at *http://206.166.205.174/*. This template represents a preliminary design that was developed as part of the Year I Tasks in the IEDR to Support the South Fork of the Broad River Research Program Scope of Work.

Users will interface with the data repository using the Internet and specially designed web pages. Users will be required to enter data at various stages of their query, download, and submittal steps. Maximum use will be made of dropdown menus and geographical map selection (i.e., search for data by clicking on map points). A general overview of the conceptual design is shown in Figure 2.

Home Page Options

Navigation throughout the repository will begin with a home page. Users visiting the data repository web site will be greeted with a home page that offers the user the following options:

a. *Obtain User ID:* A button at the home page will permit a first-time user to proceed to another page that will permit an individual to request a user ID and password. Security for the repository requires that each user obtain a valid user ID and password.

b. *Contact Information:* A list of pertinent names and numbers for contacting to obtain further information about the repository are presented.

c. *Enter Repository:* A selection on the home page menu will permit the user to proceed to the page where he/she can log onto the repository, if the user has previously obtained a user ID and password.

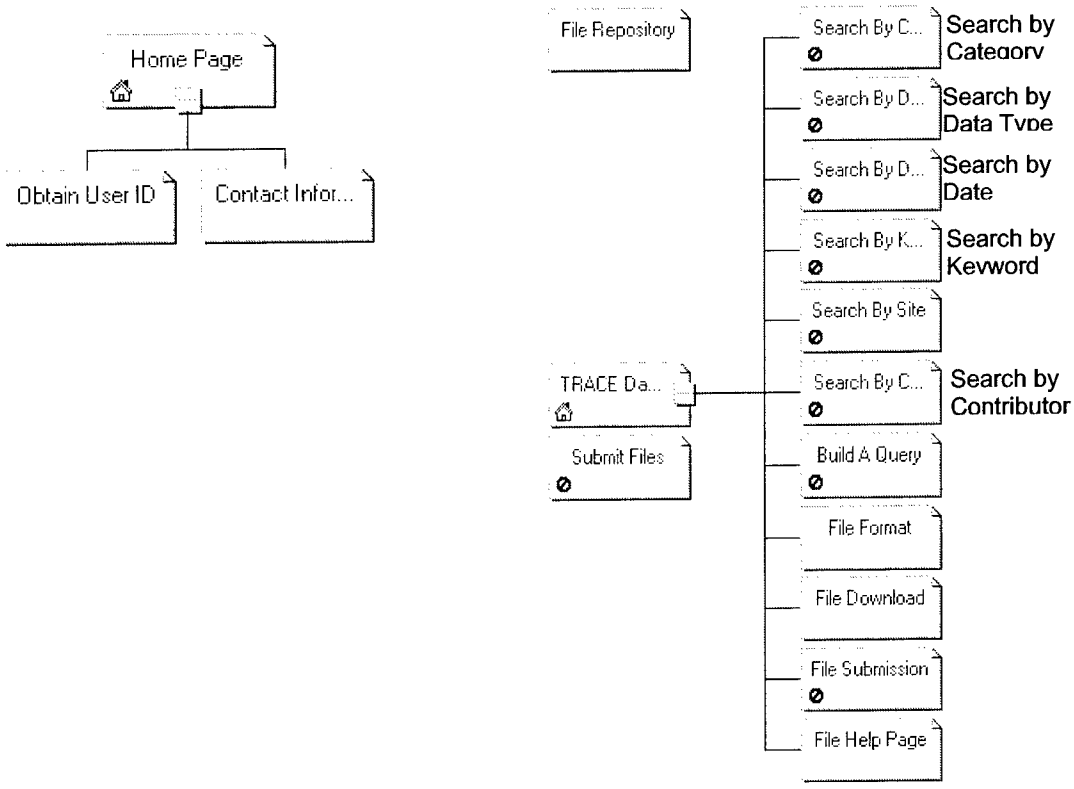


Figure 2. Conceptual Design of the IEDR at the web page level

After log on, the users shall have access to many modules within the repository.

a. *Browse data*: Users shall be permitted to select a file for downloading from a directory/subdirectory listing.

b. *Search data:* The user will be able to search the repository based on categories/themes for a listing of data files that may be selected for downloading. Several pages will be developed that permit the user to view listings of files contained in the repository based on a criterion. Users will be able to obtain file listings based on the following criterion:

(1) Categories/themes

(2) Dates

(3) Data types

(4) Contributors/authors

(5) Keywords

(6) Monitoring sites

Users may also select a page that permits them to build a query for file listings based on multiple criteria.

The above searches shall provide the user with a listing of files and amplifying data such as date, document title, metadata file, etc. Once the user had identified a file of interest, the user can select the filename and dowload the file.

c. *Submit data*: This option allows the user to submit a file to the repository.

d. *Download data*: This option provides the users with access to the file directories and subdirectories where a specific file can be selected for downloading.

e. *File format*: This page will list file formats acceptable to the repository.

f. *File Help page*: This page will include a help section that provides additional information about procedures and restrictions associated with the repository.

g. *Links*: The file repository shall provide links and compatibility with regional or other repository sites.

Preliminary web pages are presented in Appendix A.

Appendix A
Preliminary Web Pages

EPA United States Environmental Protection Agency

National Exposure Research Laboratory

Registered Users

Enter Data Repository

New Users

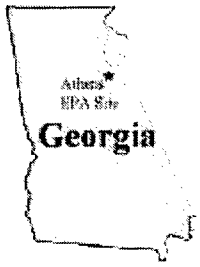
Obtain User ID

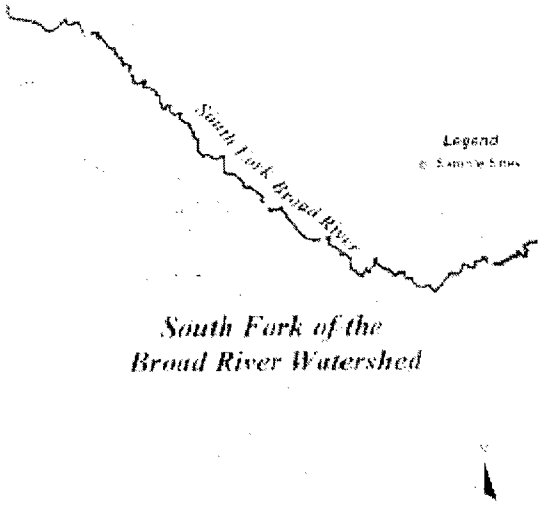
Contact Information

Technology Required for Alternative Analyses for a Changing Environment (TRACE) Program

Total Maximum Daily Load (TMDL) Field Study
South Fork of the Broad River Watershed

Welcome to the TRACE Program Site. The goal of the program is to develop and maintain a comprehensive technical support capability within NERL that directly links environmental TMDL exposure research activities and products for the EPA Office of Water and Regional Offices to be used for implementation of policy, regulatory development, remediation, and enforcement needs and activities.





South Fork of the Broad River Watershed

You are visitor:
0028
since June 22, 2001

EPA United States Environmental Protection Agency

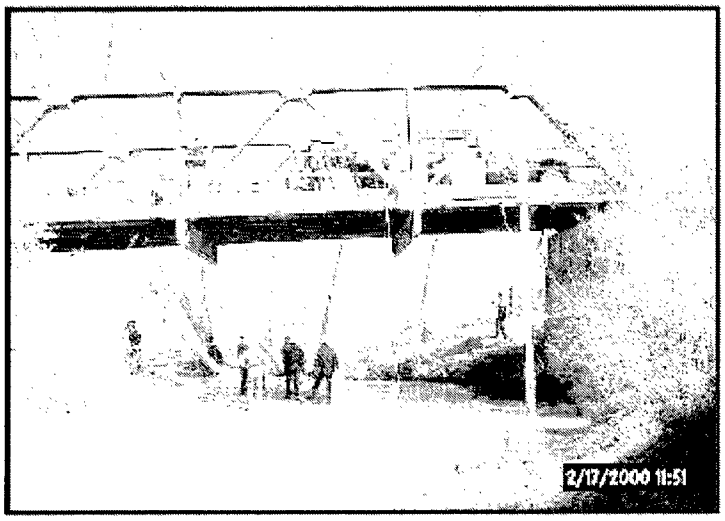
Office of Research and Development

National Exposure Research Laboratory
Ecosystems Research Division

TRACE Data Repository

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

Welcome to the National Exposure Research Laboratory (NERL) Ecosystems Research Division Integrated Electronic Data File Repository. This is a U.S. Government site to support modeling and research efforts in conjunction with risk assessment.



EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

Search By Category

Home
Search By Category
Search By Data Type
Search By Date
Search By Keyword
Search By Site
Search By Contributor
Build A Query
File Format
File Download
File Submission
File Help Page

This search provides the capability to perform file searches based on the category under which the file is stored in the repository.

Category:

Submit Query

Title	Date	Site	Data Type	Filename	Metadata

No records returned.

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

Search By Data Type

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

This area provides the capability to perform file searches based on the data type of the files in the repository

Data Type:

Submit Query

Title	Category	Date	Site	Filename	Metadata

No records retrieved

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

Search By Date

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

This area provides the capability to perform file searches based on the attributed date associated with the files in the repository

Date:

Submit Query (Use the format (MM/DD/YY))

Title	Category	Site	Data Type	Filename	Metadata

No records retrieved

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

Search By Keyword

Home
Search By Category
Search By Data Type
Search By Date
Search By Keyword
Search By Site
Search By Contributor
Build A Query
File Format
File Download
File Submission
File Help Page

This area provides the capability to perform file searches based on keywords associated with the files in the repository

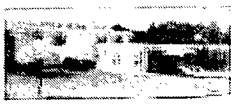
Keyword:

Submit Query

Title	Category	Date	Site	Data Type	Filename	Metadata

No records returned

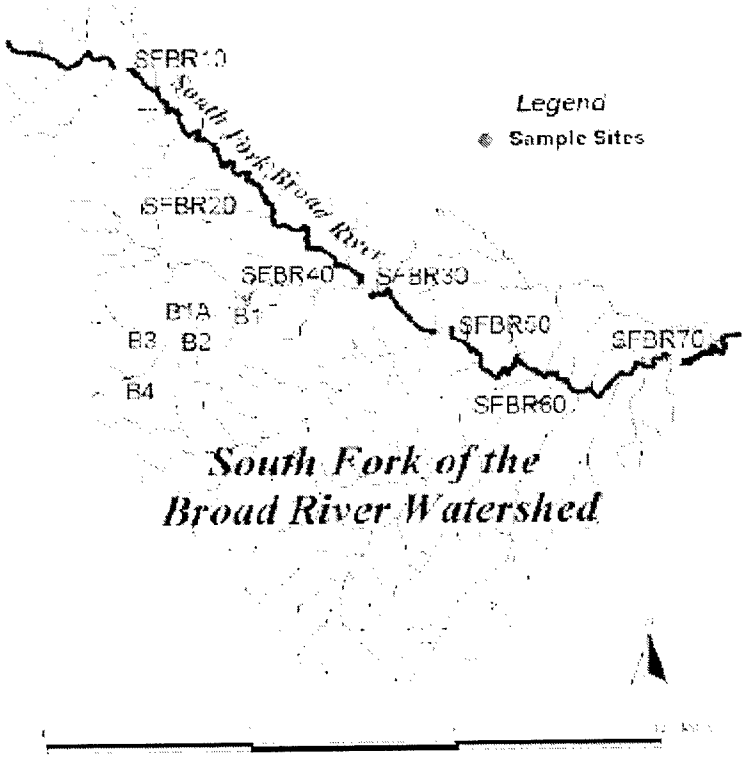
EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division



Search By Site

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

This area provides the capability to perform file searches based on the designated site associated with the files in the repository



Site:

Submit Query

Title	Category	Date	Data Type	Filename	Metadata

No records retrieved

EPA United States Environmental Protection Agency

Office of Research and Development

National Exposure Research Laboratory

Ecosystems Research Division

Search By Contributor

Home

Search By Category

Search By Date Type

Search By Date

Search By Keyword

Search By Site

Search By Contributor

Build A Query

File Format

File Download

File Submission

File Help Page

This area provides the capability to perform searches based on who contributed files to the repository

Contributor:

Submit Query

Title	Category	Date	Site	Data Type	Filename Metadata

No records returned

EPA United States Environmental Protection Agency

National Exposure Research Laboratory
Ecosystems Research Division

Build A Query

Home

Search By Category

Search By Data Type

Search By Date

Search By Keyword

Search By Site

Search By Contributor

Build A Query

File Format

File Download

File Submission

File Help Page

Build a query to find a desired file based on criteria using the 'and/or' operators.

Category:

Data Type:

Site:

Submit Query (This query is based on the 'and' operator.)

Title	Category	Date	Site	Data Type	Filename	Metadata

No records returned

Category:

Data Type:

Site:

Submit Query (This query is based on the 'or' operator.)

Title	Category	Date	Site	Data Type	Filename	Metadata

No records returned

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

File Format

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

The file format of supported files in the repository is open in the sense that it allows a multiple modeling environment. No restraints currently exist on the acceptable format of files to be uploaded to the repository.

EPA United States Environmental Protection Agency

Office of Research and Development

National Exposure Research Laboratory
Ecosystems Research Division

File Download

- Home
- Search By Category
- Search By Data Type
- Search By Date
- Search By Keyword
- Search By Site
- Search By Contributor
- Build A Query
- File Format
- File Download
- File Submission
- File Help Page

This area provides the capability to perform file downloads from the FTP portion of the repository.

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

File Submission

Home
Search By Category
Search By Data Type
Search By Date
Search By Keyword
Search By Site
Search By Contributor
Build A Query
File Format
File Download
File Submission
File Help Page

This screen provides a means to submit and catalog data. This does not provide file validation as the function rests with the contributor, author or originator of the file.

We need a record of who is contributing the data files, so please fill out the following information about yourself.

Contributor's Name:

Organization of affiliation:

City: State: AL Zip:

E-mail address:

Telephone:

Certain information is required for each file uploaded into the repository. Below, fill in all the information in the form. If you do not know the answer, please enter "N/A" where it applies.

Title:

File Name:

Metadata File Name:

What category does this data best fit into? Please Select

Please pick from the list of predefined keywords that best fit your data.

Keyword 1: Air Temperature
Keyword 2: Air Temperature
Keyword 3: Air Temperature
Keyword 4: Air Temperature

Data Date: (use this format MM/DD/YY)

Data Type:

Indicate the Site Location: Select

Give a brief description or enter comments regarding this data:

Submit Reset

EPA United States Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division

File Help Page

Home
Search By Category
Search By Data Type
Search By Date
Search By Keyword
Search By Site
Search By Contributor
Build A Query
File Format
File Download
File Submission
File Help Page

The ERD Repository is searchable by the following:

- Category
- Data Type
- Date
- Keyword
- Site
- Contributor

When a search is performed on the repository using one of these items, the results will appear listing the file, the metadata, and accompanying information associated with that file. The file and metadata allow for downloading.

The build a query page allows the user to designed their own query based upon the 'and or' operators. This allows for more manipulation of the search.

The file download page is essentially a FTP site designed for those more familiar with the repository.

The file submission page allows the user to upload files to the repository. The form must be completed so that the files can accurately be placed in the repository. This does not provide file validation as the function rests with the contributor or author of the data.

At this time, no restrictions exist on the acceptable format of files being uploaded to the repository.

206.166.205.174 - /RAD/Erd/

[To Parent Directory]
1/18/01 11:17 AM <dir> data
1/18/01 11:17 AM <dir> metadata

206.166.205.174 - /RAD/Erd/data/

[To Parent Directory]
1/18/01 11:17 AM <dir> Disharge_Data
1/18/01 11:17 AM <dir> Documentation
1/18/01 11:17 AM <dir> Flow_Data
1/18/01 11:17 AM <dir> Meteorological_Data
1/18/01 11:17 AM <dir> Nitrate_Content
1/18/01 11:17 AM <dir> Pathogen_Data
1/18/01 11:17 AM <dir> Phosphate_Content
1/18/01 11:17 AM <dir> Sediment_Data
1/18/01 11:17 AM <dir> Site_Characterization
1/18/01 11:17 AM <dir> Stream_Data

206.166.205.174 - /RAD/Erd/metadata/

[To Parent Directory]
1/18/01 11:17 AM <dir> Discharge_Data
1/18/01 11:17 AM <dir> Documentation
1/18/01 11:17 AM <dir> Flow_Data
1/18/01 11:17 AM <dir> Meteorological_Data
1/18/01 11:17 AM <dir> Nitrate_Content
1/18/01 11:17 AM <dir> Pathogen_Data
1/18/01 11:17 AM <dir> Phosphate_Content
1/18/01 11:17 AM <dir> Sediment_Data
1/18/01 11:17 AM <dir> Site_Characterization
1/18/01 11:17 AM <dir> Stream_Data

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
August 2001	Final report	

4. TITLE AND SUBTITLE

Integrated Electronic Data Repository to Support the South Fork of the Broad River Research Program

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Linda Peyman Dove, Steve Bong

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Engineer Research and Development Center
Environmental Laboratory
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Jaya Corporation
4900 University Square, Suite 30
Huntsville, AL 35816

8. PERFORMING ORGANIZATION REPORT NUMBER

ERDC/EL TR-01-14

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Environmental Protection Agency
Ecosystem Research Division, National Exposure Research Laboratory
960 College Station Road, Athens, GA 30605

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report was prepared for the Ecosystems Research Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency (USEPA), Athens, GA. It is the summary of the first part of the Integrated Electronic Data Repository (IEDR) to support all phases of the multiyear Technology Required for Alternative Analyses for a Changing Environment (TRACE) Program. The project was developed by the U.S. Army Engineer Research and Development Center (ERDC), Vicksburg, MS, to design and implement a plan for an Integrated Electronic Data Repository (IEDR) to support all phases of the multiyear Technology Required for Alternative analyses for a Changing Environment (TRACE) Program. Specifically, the repository is needed to support the Total Maximum Daily Load (TMDL) field research projects established in the South Fork of the Broad River watershed. The purpose of the repository is to archive and disseminate various types of data associated with the project to support environmental monitoring and modeling. The modeling activities include those such as sediment transport, watershed analyses, and environmental analyses. These models require data gathered ultimately from various data files submitted and cataloged in an orderly fashion with the repository.

This report summaries the design requirements and considerations used to develop a prototype web-based repository site for the TMDL research project. It is presented in two parts: design requirements and considerations and prototype web-based repository.

15. SUBJECT TERMS

Data repository
Electronic data repository
File repository
TRACE program
Web-based repository

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		26	